

OS 90

P24

#6



OIPE

ENTERED

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/10/086,816A

DATE: 11/06/2002
 TIME: 14:32:39

Input Set : A:\2401pcd.app
 Output Set: N:\CRF4\11062002\J086816A.raw

3 <110> APPLICANT: CIHLAR, TOMAS
 5 <120> TITLE OF INVENTION: NOVEL GENE ENCODING ORGANIC ANION TRANSPORTER
 7 <130> FILE REFERENCE: 240.1PCD
 9 <140> CURRENT APPLICATION NUMBER: 10/086,816A
 C--> 10 <141> CURRENT FILING DATE: 2002-10-15
 12 <150> PRIOR APPLICATION NUMBER: 09/330,245
 13 <151> PRIOR FILING DATE: 1999-06-10
 15 <150> PRIOR APPLICATION NUMBER: 60/132,267
 16 <151> PRIOR FILING DATE: 1999-05-03
 18 <150> PRIOR APPLICATION NUMBER: 60/088,864
 19 <151> PRIOR FILING DATE: 1998-06-11
 21 <160> NUMBER OF SEQ ID NOS: 10
 23 <170> SOFTWARE: PatentIn Ver. 2.1
 25 <210> SEQ ID NO: 1
 26 <211> LENGTH: 2123
 27 <212> TYPE: DNA
 28 <213> ORGANISM: Homo sapiens
 30 <220> FEATURE:
 31 <221> NAME/KEY: CDS
 32 <222> LOCATION: (263)..(1912)
 34 <400> SEQUENCE: 1
 35 ttccctgccct gaccccaaaa gtgaggagaa gctgcaaggg aaaagggagg gacagatcg 60
 37 ggagaccggg gaagaaggag gacgagccaa ggaggctgct gtccccccac agaggcagctc 120
 39 ggactcagct cccgggaagc aaccaggctg cgaggcaac ggcagtgcgt ctccctccagc 180
 41 gaaggacacgc aggccaggcag acagacagag gtcctggac tggaggcct cagccccagg 240
 43 ccactggcgt gggctggcc ca atg gcc ttt aat gac ctc ctg cag cag gtg 292
 44 Met Ala Phe Asn Asp Leu Leu Gln Gln Val
 45 1 5 10
 47 ggg ggt gtc ggc cgc ttc cag cag atc cag gtc acc ctg gtg gtc ctc 340
 48 Gly Gly Val Gly Arg Phe Gln Gln Ile Gln Val Thr Leu Val Val Leu
 49 15 20 25
 51 ccc ctg ctc ctg atg gct tct cac aac acc ctg cag aac ttc act gct 388
 52 Pro Leu Leu Met Ala Ser His Asn Thr Leu Gln Asn Phe Thr Ala
 53 30 35 40
 55 gcc atc cct acc cac cac tgc cgc ccg cct gcc gat gcc aac ctc agc 436
 56 Ala Ile Pro Thr His His Cys Arg Pro Pro Ala Asp Ala Asn Leu Ser
 57 45 50 55
 59 aag aac ggg ggg ctg gag gtc tgg ctg ccc cgg gac agg cag ggg cag 484
 60 Lys Asn Gly Gly Leu Glu Val Trp Leu Pro Arg Asp Arg Gln Gly Gln
 61 60 65 70
 63 cct gag tcc tgc ctc cgc ttc acc tcc ccg cag tgg gga ctg ccc ttt 532
 64 Pro Glu Ser Cys Leu Arg Phe Thr Ser Pro Gln Trp Gly Leu Pro Phe
 65 75 80 85 90

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/086,816A

DATE: 11/06/2002

TIME: 14:32:39

Input Set : A:\2401pcd.app

Output Set: N:\CRF4\11062002\J086816A.raw

67	ctc	aat	ggc	aca	gaa	gcc	aat	ggc	aca	ggg	gcc	aca	gag	ccc	tgc	acc	580
68	Leu	Asn	Gly	Thr	Glu	Ala	Asn	Gly	Thr	Gly	Ala	Thr	Glu	Pro	Cys	Thr	
69				95					100							105	
71	gat	ggc	tgg	atc	tat	gac	aac	agc	acc	tcc	cca	tct	acc	atc	gtg	act	628
72	Asp	Gly	Trp	Ile	Tyr	Asp	Asn	Ser	Thr	Phe	Pro	Ser	Thr	Ile	Val	Thr	
73				110					115							120	
75	gag	tgg	gac	ctt	gtg	tgc	tct	cac	agg	gcc	cta	cgc	cag	ctg	gcc	cag	676
76	Glu	Trp	Asp	Leu	Val	Cys	Ser	His	Arg	Ala	Leu	Arg	Gln	Leu	Ala	Gln	
77				125				130								135	
79	tcc	ttg	tac	atg	gtg	ggg	gtg	ctg	ctc	gga	gcc	atg	gtg	ttc	ggc	tac	724
80	Ser	Leu	Tyr	Met	Val	Gly	Val	Leu	Leu	Gly	Ala	Met	Val	Phe	Gly	Tyr	
81				140				145								150	
83	ctt	gca	gac	agg	cta	ggc	cgc	cgg	aag	gta	ctc	atc	ttg	aac	tac	ctg	772
84	Leu	Ala	Asp	Arg	Leu	Gly	Arg	Arg	Lys	Val	Leu	Ile	Leu	Asn	Tyr	Leu	
85	155				160					165						170	
87	cag	aca	gct	gtg	tca	ggg	acc	tgc	gca	gcc	ttc	gca	ccc	aac	tcc	ccc	820
88	Gln	Thr	Ala	Val	Ser	Gly	Thr	Cys	Ala	Ala	Phe	Ala	Pro	Asn	Phe	Pro	
89				175				180								185	
91	atc	tac	tgc	gcc	ttc	cgg	ctc	tcg	ggc	atg	gtc	gct	gtc	ggc	atc		868
92	Ile	Tyr	Cys	Ala	Phe	Arg	Leu	Leu	Ser	Gly	Met	Ala	Leu	Ala	Gly	Ile	
93				190				195								200	
95	tcc	ctc	aac	tgc	atg	aca	ctg	aat	gtg	gag	tgg	atg	ccc	att	cac	aca	916
96	Ser	Leu	Asn	Cys	Met	Thr	Leu	Asn	Val	Glu	Trp	Met	Pro	Ile	His	Thr	
97				205				210								215	
99	cgg	gcc	tgc	gtg	ggc	acc	ttg	att	ggc	tat	gtc	tac	agc	ctg	ggc	cag	964
100	Arg	Ala	Cys	Val	Gly	Thr	Leu	Ile	Gly	Tyr	Val	Tyr	Ser	Leu	Gly	Gln	
101				220				225								230	
103	tcc	ctc	ctg	gct	ggt	gtg	gcc	tac	gct	gtg	ccc	cac	tgg	cgc	cac	ctg	1012
104	Phe	Leu	Leu	Ala	Gly	Val	Ala	Tyr	Ala	Val	Pro	His	Trp	Arg	His	Leu	
105	235				240				245							250	
107	cag	cta	ctg	gtc	tct	gcg	cct	ttt	ttt	gcc	ttc	ttc	atc	tac	tcc	tgg	1060
108	Gln	Leu	Leu	Val	Ser	Ala	Pro	Phe	Phe	Ala	Phe	Phe	Ile	Tyr	Ser	Trp	
109				255				260								265	
111	ttc	ttc	att	gag	tcg	gcc	cgc	tgg	cac	tcc	tcc	ggg	agg	ctg	gac		1108
112	Phe	Phe	Ile	Glu	Ser	Ala	Arg	Trp	His	Ser	Ser	Ser	Gly	Arg	Leu	Asp	
113				270				275								280	
115	ctc	acc	ctg	agg	gcc	ctg	cag	aga	gtc	gcc	cgg	atc	aat	ggg	aag	cgg	1156
116	Leu	Thr	Leu	Arg	Ala	Leu	Gln	Arg	Val	Ala	Arg	Ile	Asn	Gly	Lys	Arg	
117				285				290								295	
119	gaa	gaa	gga	gcc	aaa	ttg	agt	atg	gag	gta	ctc	cgg	gcc	agt	ctg	cag	1204
120	Glu	Glu	Gly	Ala	Lys	Leu	Ser	Met	Glu	Val	Leu	Arg	Ala	Ser	Leu	Gln	
121				300				305								310	
123	aag	gag	ctg	acc	atg	ggc	aaa	ggc	cag	gca	tcg	gcc	atg	gag	ctg	ctg	1252
124	Lys	Glu	Leu	Thr	Met	Gly	Lys	Gly	Gln	Ala	Ser	Ala	Met	Glu	Leu	Leu	
125	315				320				325							330	
127	cgc	tgc	ccc	acc	ctc	cgc	cac	ctc	ttc	ctc	tgc	ctc	tcc	atg	ctg	tgg	1300
128	Arg	Cys	Pro	Thr	Leu	Arg	His	Leu	Phe	Leu	Cys	Leu	Ser	Met	Leu	Trp	
129				335				340								345	
131	ttt	gcc	act	agc	ttt	gca	tac	tat	ggg	ctg	gtc	atg	gac	ctg	cag	ggc	1348

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/086,816A

DATE: 11/06/2002

TIME: 14:32:40

Input Set : A:\2401pcd.app

Output Set: N:\CRF4\11062002\J086816A.raw

132	Phe	Ala	Thr	Ser	Phe	Ala	Tyr	Tyr	Gly	Leu	Val	Met	Asp	Leu	Gln	Gly	
133					350				355					360			
135	ttt	gga	gtc	agc	atc	tac	cta	atc	cag	gtg	atc	ttt	ggt	gct	gtg	gac	1396
136	Phe	Gly	Val	Ser	Ile	Tyr	Leu	Ile	Gln	Val	Ile	Phe	Gly	Ala	Val	Asp	
137					365				370				375				
139	ctg	cct	gcc	aag	ctt	gtg	ggc	ttc	ctt	gtc	atc	aac	tcc	ctg	ggt	cgc	1444
140	Leu	Pro	Ala	Lys	Leu	Val	Gly	Phe	Leu	Val	Ile	Asn	Ser	Leu	Gly	Arg	
141					380				385				390				
143	cgg	cct	gcc	cag	atg	gct	gca	ctg	ctg	gca	ggc	atc	tgc	atc	ctg		1492
144	Arg	Pro	Ala	Gln	Met	Ala	Ala	Leu	Leu	Leu	Ala	Gly	Ile	Cys	Ile	Leu	
145					395				400			405		410			
147	ctc	aat	ggg	gtg	ata	ccc	cag	gac	cag	tcc	att	gtc	cga	acc	tct	ctt	1540
148	Leu	Asn	Gly	Val	Ile	Pro	Gln	Asp	Gln	Ser	Ile	Val	Arg	Thr	Ser	Leu	
149					415				420			425					
151	gct	gtg	ctg	ggg	aag	ggt	tgt	ctg	gct	gcc	tcc	ttc	aac	tgc	atc	ttc	1588
152	Ala	Val	Leu	Gly	Lys	Gly	Cys	Leu	Ala	Ala	Ser	Phe	Asn	Cys	Ile	Phe	
153					430				435			440					
155	ctg	tat	act	ggg	gaa	ctg	tat	ccc	aca	atg	atc	cgg	cag	aca	ggc	atg	1636
156	Leu	Tyr	Thr	Gly	Glu	Leu	Tyr	Pro	Thr	Met	Ile	Arg	Gln	Thr	Gly	Met	
157					445				450			455					
159	gga	atg	ggc	agc	acc	atg	gcc	cga	gtg	ggc	agc	atc	gtg	agc	cca	ctg	1684
160	Gly	Met	Gly	Ser	Thr	Met	Ala	Arg	Val	Gly	Ser	Ile	Val	Ser	Pro	Leu	
161					460				465			470					
163	gtg	agc	atg	act	gcc	gag	ctc	tac	ccc	tcc	atg	cct	ctc	ttc	atc	tac	1732
164	Val	Ser	Met	Thr	Ala	Glu	Leu	Tyr	Pro	Ser	Met	Pro	Leu	Phe	Ile	Tyr	
165					475				480			485		490			
167	ggt	gct	gtt	cct	gtg	gcc	gcc	agc	gct	gtc	act	gtc	ctc	ctg	cca	gag	1780
168	Gly	Ala	Val	Pro	Val	Ala	Ala	Ser	Ala	Val	Thr	Val	Leu	Leu	Pro	Glu	
169					495				500			505					
171	acc	ctg	ggc	cag	cca	ctg	cca	gac	acg	gtg	cag	gac	ctg	gag	agc	agg	1828
172	Thr	Leu	Gly	Gln	Pro	Leu	Pro	Asp	Thr	Val	Gln	Asp	Leu	Glu	Ser	Arg	
173					510				515			520					
175	aaa	ggg	aaa	cag	acg	cga	cag	caa	caa	gag	cac	cag	aag	tat	atg	gtc	1876
176	Lys	Gly	Lys	Gln	Thr	Arg	Gln	Gln	Gln	Glu	His	Gln	Lys	Tyr	Met	Val	
177					525				530			535					
179	cca	ctg	cag	gcc	tca	gca	caa	gag	aag	aat	gga	ctc	tgaggactga				1922
180	Pro	Leu	Gln	Ala	Ser	Ala	Gln	Glu	Lys	Asn	Gly	Leu					
181					540				545			550					
183	gaaggggcct	tacagaaccc	taaaggggagg	gaaggtccta	caggtctccg	gccacccaca											1982
185	caaggaggag	gaagaggaaa	tttgtgacca	agtgtggggg	tttgtgggtca	ggaaagcatc											2042
187	ttccccagggg	tccacacctcc	tttataaacc	ccaccagaac	cacatcatta	aaaggtttga											2102
189	ctgcgaaaaaa	aaaaaaaaaa	a														2123
192	<210>	SEQ ID NO:	2														
193	<211>	LENGTH:	550														
194	<212>	TYPE:	PRT														
195	<213>	ORGANISM:	Homo sapiens														
197	<400>	SEQUENCE:	2														
198	Met	Ala	Phe	Asn	Asp	Leu	Leu	Gln	Gln	Val	Gly	Gly	Val	Gly	Arg	Phe	
199	1				5					10				15			

RAW SEQUENCE LISTING DATE: 11/06/2002
 PATENT APPLICATION: US/10/086,816A TIME: 14:32:40

Input Set : A:\2401pcd.app
 Output Set: N:\CRF4\11062002\J086816A.raw

201 Gln Gln Ile Gln Val Thr Leu Val Val Leu Pro Leu Leu Met Ala
 202 20 25 30
 204 Ser His Asn Thr Leu Gln Asn Phe Thr Ala Ala Ile Pro Thr His His
 205 35 40 45
 207 Cys Arg Pro Pro Ala Asp Ala Asn Leu Ser Lys Asn Gly Gly Leu Glu
 208 50 55 60
 210 Val Trp Leu Pro Arg Asp Arg Gln Gly Gln Pro Glu Ser Cys Leu Arg
 211 65 70 75 80
 213 Phe Thr Ser Pro Gln Trp Gly Leu Pro Phe Leu Asn Gly Thr Glu Ala
 214 85 90 95
 216 Asn Gly Thr Gly Ala Thr Glu Pro Cys Thr Asp Gly Trp Ile Tyr Asp
 217 100 105 110
 219 Asn Ser Thr Phe Pro Ser Thr Ile Val Thr Glu Trp Asp Leu Val Cys
 220 115 120 125
 222 Ser His Arg Ala Leu Arg Gln Leu Ala Gln Ser Leu Tyr Met Val Gly
 223 130 135 140
 225 Val Leu Leu Gly Ala Met Val Phe Gly Tyr Leu Ala Asp Arg Leu Gly
 226 145 150 155 160
 228 Arg Arg Lys Val Leu Ile Leu Asn Tyr Leu Gln Thr Ala Val Ser Gly
 229 165 170 175
 231 Thr Cys Ala Ala Phe Ala Pro Asn Phe Pro Ile Tyr Cys Ala Phe Arg
 232 180 185 190
 234 Leu Leu Ser Gly Met Ala Leu Ala Gly Ile Ser Leu Asn Cys Met Thr
 235 195 200 205
 237 Leu Asn Val Glu Trp Met Pro Ile His Thr Arg Ala Cys Val Gly Thr
 238 210 215 220
 240 Leu Ile Gly Tyr Val Tyr Ser Leu Gly Gln Phe Leu Leu Ala Gly Val
 241 225 230 235 240
 243 Ala Tyr Ala Val Pro His Trp Arg His Leu Gln Leu Leu Val Ser Ala
 244 245 250 255
 246 Pro Phe Phe Ala Phe Phe Ile Tyr Ser Trp Phe Phe Ile Glu Ser Ala
 247 260 265 270
 249 Arg Trp His Ser Ser Ser Gly Arg Leu Asp Leu Thr Leu Arg Ala Leu
 250 275 280 285
 252 Gln Arg Val Ala Arg Ile Asn Gly Lys Arg Glu Glu Gly Ala Lys Leu
 253 290 295 300
 255 Ser Met Glu Val Leu Arg Ala Ser Leu Gln Lys Glu Leu Thr Met Gly
 256 305 310 315 320
 258 Lys Gly Gln Ala Ser Ala Met Glu Leu Leu Arg Cys Pro Thr Leu Arg
 259 325 330 335
 261 His Leu Phe Leu Cys Leu Ser Met Leu Trp Phe Ala Thr Ser Phe Ala
 262 340 345 350
 264 Tyr Tyr Gly Leu Val Met Asp Leu Gln Gly Phe Gly Val Ser Ile Tyr
 265 355 360 365
 267 Leu Ile Gln Val Ile Phe Gly Ala Val Asp Leu Pro Ala Lys Leu Val
 268 370 375 380
 270 Gly Phe Leu Val Ile Asn Ser Leu Gly Arg Arg Pro Ala Gln Met Ala
 271 385 390 395 400
 273 Ala Leu Leu Leu Ala Gly Ile Cys Ile Leu Leu Asn Gly Val Ile Pro

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/086,816A

DATE: 11/06/2002

TIME: 14:32:40

Input Set : A:\2401pcd.app
 Output Set: N:\CRF4\11062002\J086816A.raw

```

274          405          410          415
276 Gln Asp Gln Ser Ile Val Arg Thr Ser Leu Ala Val Leu Gly Lys Gly
277          420          425          430
279 Cys Leu Ala Ala Ser Phe Asn Cys Ile Phe Leu Tyr Thr Gly Glu Leu
280          435          440          445
282 Tyr Pro Thr Met Ile Arg Gln Thr Gly Met Gly Met Gly Ser Thr Met
283          450          455          460
285 Ala Arg Val Gly Ser Ile Val Ser Pro Leu Val Ser Met Thr Ala Glu
286 465          470          475          480
288 Leu Tyr Pro Ser Met Pro Leu Phe Ile Tyr Gly Ala Val Pro Val Ala
289          485          490          495
291 Ala Ser Ala Val Thr Val Leu Leu Pro Glu Thr Leu Gly Gln Pro Leu
292          500          505          510
294 Pro Asp Thr Val Gln Asp Leu Glu Ser Arg Lys Gly Lys Gln Thr Arg
295          515          520          525
297 Gln Gln Gln Glu His Gln Lys Tyr Met Val Pro Leu Gln Ala Ser Ala
298          530          535          540
300 Gln Glu Lys Asn Gly Leu
301 545          550
304 <210> SEQ ID NO: 3
305 <211> LENGTH: 77
306 <212> TYPE: DNA
307 <213> ORGANISM: Artificial Sequence
309 <220> FEATURE:
310 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
311     oligonucleotide
313 <400> SEQUENCE: 3
314 accgtctaga attctttta ttttaattt tctttcaaattt acgtccacca tggcctttaa 60
315 tgaccccttg cagcagg          77
318 <210> SEQ ID NO: 4
319 <211> LENGTH: 51
320 <212> TYPE: DNA
321 <213> ORGANISM: Artificial Sequence
323 <220> FEATURE:
324 <223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic
325     oligonucleotide
327 <400> SEQUENCE: 4
328 tactcacgtg gatcctgatc agacgtctgt aggaccttcc ctcccttag g          51
331 <210> SEQ ID NO: 5
332 <211> LENGTH: 9
333 <212> TYPE: DNA
334 <213> ORGANISM: Artificial Sequence
336 <220> FEATURE:
337 <223> OTHER INFORMATION: Description of Artificial Sequence: Kozak
338     consensus sequence
340 <400> SEQUENCE: 5
341 ccaccatgg          9
344 <210> SEQ ID NO: 6
345 <211> LENGTH: 25

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/086,816A

DATE: 11/06/2002

TIME: 14:32:41

Input Set : A:\2401pcd.app

Output Set: N:\CRF4\11062002\J086816A.raw

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date